



Fig2. The operation process of the improved Apriori algorithm

Conclusion

The biggest drawback of traditional Apriori algorithm is that each item in the candidates of frequent item sets must scan database one time, and the efficiency of execution is very low. we proposed an improvement, this algorithm scan the transaction database D only when L₁-candidates is produced, the rest scan the preceding result instead of the transaction database, greatly reducing I/O load and increase efficiency. The interesting association rule can be accessed more effectively by the improved Apriori algorithm.

References

- [1] Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth. From Data Mining to Knowledge Discovery in Database [J]. American Association for Artificial Intelligence, AAAI Press, July, 1996, 37-54.
- [2] M.S.Chen, J.Han, and P.S. Yu. Data Mining: An Overview from Database Perspective [J]. Knowledge and Data Engineering, 1996 (8) 866-883.
- [3] A.Chen, N.Chen, L.X.Zhou. The Technology and Appliace of Data Mining[M]. BeiJing: Science Press 2006.
- [4] W.W. Chen. Data Warehouse and Data Mining[M]. BeiJing: Tsinghua University Press 2006.
- [5] J. Cheng, Y.P. Ke, Wilfred Ng. Effective Elimination of Redundant Association Rules [J]. Data Mining and Knowledge Discovery. 2008 (16) 221-249
- [6] R.Agrawal, and R.Srikant. Fast Apriorithms for Mining Assiciation Rules [A]. Proc. the 20th VLDB[C]. 1994 487-499.
- [7] Brin S, Motwai R, Ullman J D, Tsur S. Dynamic Itemset Counting and Implication Rules for market basket data[A]. Proceedings of the ACM SIGMOD International Conference on Management of Data[C].1997 255-264
- [8] J.S.Park, M.S. Chen, and P.S. Yu. An Effective Hash Based Algorithm for Mining Association Rules[A]. Proc.ACM SGMOD[C]. 1995 175-185.
- [9] H.Shen, B.L.Lu, M.S. Utiyama. Comparison and Improvements of Feature Extraction Methods for Text Categorization [J]. Computer Simulation, 2006 (23) 222-224.
- [10] L.L Dai, H.Y. Huang, and Z.X. Chen. A Comparative Study on Feature Selection in Chinese Text Categorization [J]. Journal of Chinese Information Processing , 2004(18) 26-32.
- [11] Q. Zhou, M.S. ZHAO. Study on Feature Selection in Chinese Text Categorization [J]. Journal of Chinese Information Processing, 2004 (18) 17-23.
- [12] H.X Chai, Y.Wang. Improvement of Apriori Algorithm[J]. Computer Engineering and Applications, 2007 (43) 158-161.